Abstract number
1742577

Ontology-based Semantic Similarity Measures in Drug Safety: A Literature Review

François Haguinet[1], Jeffery L. Painter[2], Andrew Bate[3,4]

[1]GSK, Wavre, Belgium; [2]GSK, Durham, NC, USA; [3]GSK, Brentford, Middlesex, UK; [4]London School of Hygiene and Tropical Medicine, London, UK

Disclosures: All authors are employed by GSK and hold financial equities in GSK. All authors declare no other financial or non-financial interests, relationships or activities.

Conclusions

There are only few available applications published, primarily focusing on clustering of medical concepts.

The evaluation of various ontology-based SSMs remains sparse.

These measures hold promise for applications in drug safety, including outcome definition and information retrieval from diverse data sources.

Background

- MedDRA PTs are often used in pharmacovigilance (PV) for reporting of adverse events;[1] yet individual PTs may not adequately capture related events within the same medical context.[2]
- Customized groupings by medical experts may oversimplify the relationship between PTs and provide only a binary outcome.
- Ontology-based semantic similarity measures (SSMs) offer a potential solution to the challenge of defining outcomes, especially when retrieving information from multiple safety data sources.
- To our knowledge there is no literature overview of the use of SSMs in PV.

Aims

Explore how ontology-based SSMs have been applied in PV and the broader biomedical field.

Identify opportunities for their use in PV.

Methods

Literature review conducted in **Embase**

05-February-2024

English language

Gene-related studies were excluded*

*Only for the refined Search 1

Initial search

"semantic similarity" *or* "semantic distance"

Initial search was followed by subsequent refinements**

**Additional restrictions to publications types "Article" and "Review" were also made.

Results

Initial search

(including "semantic similarity" *or* "semantic distance")

1,503 papers

Refined search

Search 1 → 35 papers
(including "ontology" *and* "medical concepts"; excluding "gene-related")

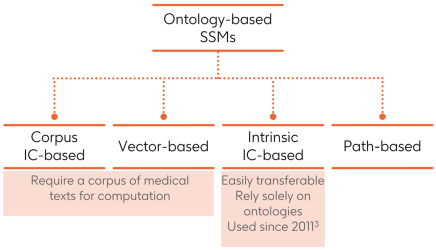
Search 2 → 14 papers
(including "safety" *or* "pharmacovigilance")

Search 3 → 5 papers
(including "MedDRA" *or* "WHO-ART")

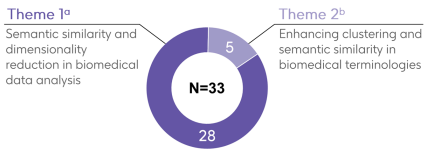
Search 4 → 50 papers
(including [Initial search *and* Search 1] *or* [Initial search *and* Search 2] *or* [Initial search *and* Search 3])

Included → 33 papers

- Ontology-based SSMs were categorized into four main types.



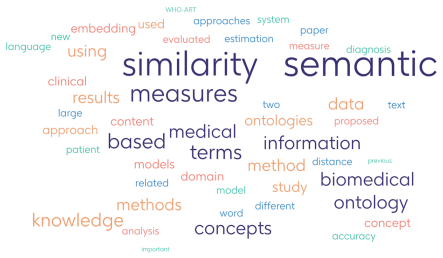
- The included papers were grouped into two themes



[a]Focus on methods to enhance the analysis, integration, and understanding of biomedical data by using SSMs and dimensionality reduction techniques to improve the machine's ability to interpret complex medical data, facilitate data exploration, and support both clinical and translational research.

[b]Focus on methods to improve clustering and SSMs in biomedical terminologies through advanced techniques like machine learning, ontological knowledge, and deep learning, aiming to increase the accuracy and utility of biomedical data for better signal detection, data mining, and knowledge discovery.

- Refinements based on specific terms yielded relevant publications, although some focused on gene ontology or were unrelated to PV.



The most used ontologies: MeSH | SNOMED-CT | UMLS

- Evaluations typically utilized the reference set of similar medical terms,[4] although some studies employed larger sets.
- Few evaluations of SSM types beyond intrinsic IC-based SSMs were found, and while some showed promising performance, they were not conclusive, especially within the realm of drug safety.

Abbreviations

MedDRA, Medical Dictionary for Regulatory Activities; PTs, Preferred Terms; WHO-ART, World Health Organization Adverse Reaction Terminology; IC, information content; N, number of analyzed papers; MeSH, Medical Subject Headings; SNOMED-CT, Systematized Nomenclature of Medicine-Clinical Terms; UMLS, Unified Medical Language System.

References

1. Dutta A. https://meddra.org/sites/default/files/page/documents_insert/meddra_-_terminologies_coding.pdf
2. Hauben M, Bate A. *Drug Discov Today*. 2009;14(7-8):343–357
3. Sánchez D, Batet M. *J Biomed Inform*. 2011;44(5):749–759
4. Pedersen T, et al. *J Biomed Inform*. 2007;40(3):288–299

Acknowledgements

Medical writing (Sara-Teodora Vulcu), design and coordination support were provided by Akkodis Belgium c/o GSK

Funding

GSK

Digital poster
Supplemental data

ISPE | 24–28 August 2024 | Berlin, Germany

Presenting author: François Haguinet, francois.f.haguinet@gsk.com

GSK